

SUR LES CONDITIONS QUI PERMETTENT D'UTILISER LES
MATRICES RUSSELLIENNES DES ANTINOMIES (1905) POUR
EXPRIMER LES THÉORÈMES DE LIMITATIONS INTERNES
DES FORMALISMES

JULES VUILLEMIN

§1. *Simplification de la matrice de l'antinomie. Forme de cette matrice chez Carnap.*

Peut-on simplifier la matrice de l'antinomie, telle que Russell la formule en 1905? On considèrera successivement les deux cas des antinomies logiques et des antinomies sémantiques.

I. La forme (A) des antinomies logiques peut être réduite à la forme (B), des qu'on abandonne l'idée de solutions alternatives¹. De plus, la forme (A) correspond aux antinomies proprement mathématiques (paradoxe de Cantor et paradoxe de Burali-Forti), tandis la forme (B) convient à l'expression des antinomies proprement logiques. Les complications qu'introduit en (A) la considération du foncteur $f'u$ sont dues à la nécessité de spécifier - pour dériver la contradiction - que l'ensemble des parties d'ensemble est un ensemble et n'est pas élément de cet ensemble ou que le nombre ordinal d'un ensemble est successeur de l'ensemble. Or la genèse de la contradiction est indépendante de ces précisions et dépend uniquement, comme le prouve la forme simplifiée (B), des propriétés logiques de la matrice. On pourra donc se limiter à l'étude de cette forme:

$$(B) (u) \{ [(x)(x\epsilon u \supset x \notin x)] \supset (u \notin u) \} \supset \\ \{ [(\exists w)(y)(y \notin y \equiv y \epsilon w)] \supset [(\exists w)(w \notin w \cdot w \epsilon w)] \}.$$

La première réduction qu'on peut proposer a trait au premier membre de l'implication. En effet, le second membre:

1. Vuillemin, *L'origine et le mécanisme des antinomies dans la première philosophie de Russell (1903)*, in *Logique et Analyse*, avril 1964, 25-26, p. 59-60. Dans cet article, on a formulé la matrice des antinomies sous la forme générale:

$$(A) (u) \{ [(x)(x\epsilon u \supset \phi x)] \supset [\phi'f'u \cdot f'u \notin u] \} \supset \\ \{ [(\exists w)(y)(\phi y \equiv y \epsilon w)] \supset [(\exists w)(\phi f'w \cdot \sim \phi f'w)] \}.$$

$$(B_1) \quad [(\exists w)(y)(y \notin y \equiv y \in w)] \supset [(\exists w)(w \notin w \cdot w \in w)]$$

est une tautologie, qui ne dépend que de la substitution de w à y , substitution légitime étant donnée la liaison universelle de y . La stipulation du premier membre de $[B]$, soit:

$$(B_2) \quad (u) [(x)(x \in u \supset x \notin x)] \supset (u \notin u)$$

n'était, en réalité, nécessaire que sous la forme complète (A), dans laquelle elle précisait les conditions dans lesquelles un ensemble doit être tenu pour n'étant pas élément d'un autre ensemble. Elle retenait d'ailleurs une condition caractéristique des raisonnements "diagonaux", dont on vérifiera aisément la présence dans les formes (A_c) ou (B_c) par lesquelles on peut exprimer le théorème de Cantor².

Sous la forme simplifiée (B₁), l'antinomie logique s'exprime, pour ainsi dire, plus librement. Soit, par exemple, l'antinomie russellienne de l'ensemble de tous les ensembles qui ne sont pas éléments d'eux-mêmes. On justifiera le caractère analytique de l'implication en (B₁) par un dilemme: ou bien cet ensemble est élément de lui-même, mais par l'identité du membre gauche de l'implication en (B₁) il n'est pas élément de lui-même; ou bien il n'est pas élément de lui-même et, par cette même identité, il est élément de lui-même. Si nous rétablissons le raisonnement qu'exprime la matrice complète (B), on dira, au contraire: comme, quel que soit l'ensemble y , si cet ensemble appartient à w , il ne se contient pas lui-même comme élément (implication déduite de l'identité du membre gauche de (B₁)), et que, par conséquent, en vertu de l'implication résultant du membre droit de (B₂), w ne peut pas ne pas être élément de lui-même, w se contient lui-même comme élément; ensuite, de cette supposition, on tire la contradiction par le même raisonnement que précédemment³.

2. *Viz op. cit.*, pp. 76-77

$$(A_c) \quad (A) \{[(x)(x \in A \supset x \notin A_x)] \supset (x_A \notin A)\} \supset \\ \{[(\exists L)(y \notin A_y \equiv y \in L)] \supset [(\exists L)(y_L \notin L \cdot y_L \in L)]\}$$

$$(B_c) \quad (A) \{[(x)(x \in A \supset x \notin x)] \supset [A \notin A]\} \supset \\ \{[(\exists L)(y)(y \notin y \equiv y \in L)] \supset [(\exists L)(L \notin L \cdot L \in L)]\}$$

3. Kleene (*Introduction to Metamathematics*, 1962, p. 37) analyse ainsi ce paradoxe: Soit T l'ensemble de tous les ensembles qui ne sont pas éléments d'eux-mêmes. "T est-il élément de lui-même? Supposons que T est un élément de lui-même, *id est* symboliquement $T \in T$. La supposition dit que T est un élément de T , *id est* T est un élément de l'ensemble de tous les ensembles qui ne sont pas éléments d'eux-mêmes, c'est-à-dire que T est un ensemble qui n'est pas élément de lui-même, *id est* symboliquement $T \notin T$. Ceci contredit la supposition $T \in T$. Jusque'ici nous n'avons pas de paradoxe, puisque la contradiction entre $T \in T$ et $T \notin T$ ne s'est produite que sous la supposition $T \in T$. Par la réduction à l'absurde, nous concluons que la supposition est fautive. Ainsi nous avons maintenant prouvé entièrement, sans supposition, que $T \notin T$. A partir du résultat établi $T \notin T$, nous pouvons continuer l'argumentation. Le résultat dit que T n'est pas élément de l'ensemble de tous les ensembles qui ne sont pas éléments d'eux-mêmes, *id est* T est un ensemble qui n'est pas élément de lui-même, *id est* symboliquement $T \in T$. A présent $T \notin T$ et $T \in T$ ont été tous deux établis et nous avons un paradoxe". La première partie du raisonnement - l'établissement du résultat $T \notin T$ - correspond à la partie (B₂) de la matrice (B), la seconde partie - l'établissement du résultat $T \in T$ - à la partie (B₁).

La forme simplifiée et dilemmatique (B₁) contient des quantificateurs et elle exprime des implications entre des rapports d'appartenance d'éléments à des classes. Le paradoxe russellien de l'imprédictable prouve qu'on peut changer cette forme en la forme:

$$(B_3) \quad [\sim F(F) \equiv \text{Impr}(F)] \supset [\sim \text{Impr}(\text{Impr}) \equiv \text{Impr}(\text{Impr})]$$

de laquelle on a éliminé les quantificateurs et où l'on a substitué des rapports de prédicats aux rapports d'appartenance d'éléments aux classes⁴. Toutefois, l'élimination du quantificateur universel est obtenue uniquement en ce que (B₃) est une proposition "ouverte" - contenant une variable libre *F*, qu'on interprêtera comme un "prédicat quelconque", tandis que (B₁) est une proposition "fermée", où figure le mot "tous". Quant à l'élimination du quantificateur existentiel, elle est apparente, puisqu'on substitue une constante: "imprédictable" à la liaison existentielle ($\exists w$) (. . . $y \varepsilon w$). D'autre part, comme on peut faire correspondre à toute classe son prédicat définissant, (B₃) correspond exactement à (B₁), à cette différence près qu'on dispose d'une proposition ouverte et non fermée et qu'on remplace l'existentiel par une constante (et que dans l'impliqué on substitue une équivalence à une conjonction). On peut écrire (B₃) sous la forme

$$(B_4) \quad [(F) F \not\equiv F \varepsilon I] \supset [I \not\equiv I \cdot I \varepsilon I],$$

où *I* est la classe qui correspond au prédicat définissant Impr. Peut-on compléter la ressemblance et écrire (B₄) sous une forme complètement quantifiée (B₅) et donc formellement identique à (B₁)

$$(B_5) \quad [(\exists G)(F) (F \not\equiv F \varepsilon G)] \supset [(\exists G)(G \not\equiv G \cdot G \varepsilon G)] ?$$

Cette transformation est entièrement légitime. Elle serait même nécessaire si nous ne disposions pas de nom propre - tel qu' *I* - pour désigner la classe associée au prédicat Impr. Donc (B₁) et (B₃) sont analogues et (B₁) offre l'avantage de faire l'économie de noms propres dans la langue formelle.

II. En 1903, Russell ne formulait qu'une antinomie sémantique, dérivée de la forme générale (A), soit⁵:

$$(A_p) \quad (u) \{ [(x)(x \varepsilon u \supset [(\exists q_x)(x \leftrightarrow q_x \cdot x \not\equiv q_x)]) \supset p_u \not\equiv u] \} \supset \\ \{ [(\exists w)(y)[(\exists q_y)(y \leftrightarrow q_y \cdot y \not\equiv q_y)] \equiv y \varepsilon w] \} \supset \\ [(\exists w)(p_w \not\equiv w \cdot p_w \varepsilon w)] \}.$$

4. C'est la forme de l'antinomie telle qu'elle apparaît dans R. Carnap, (*The Logical Syntax of Language*, transl. A. Smeaton, Routledge et Kegan, London, 1954, §60a, p. 211).

5. Vuillemin, *op. cit.*, p. 84. J'ai omis, dans l'écriture de cette matrice une partie de l'impliqué dans le premier membre de l'implication. En effet, en vertu de l'assignation (1) appliquée à la matrice (A) et donnée ici même, p. 4, on devrait avoir:

$$[(\exists q_{p_u})(p_u \leftrightarrow q_{p_u} \cdot p_u \not\equiv q_{p_u})] \cdot p_u \not\equiv u.$$

Mais, en vertu de l'univocité, $q_{p_u} = u$ et comme on a posé que $p_u \leftrightarrow u$ signifie $p = p_u$, la simplification obtenue a lieu - Il en va de même pour le conséquent dans le second membre (voir *op. cit.*, note 31).

Cette formule est interprétable dans l'antinomie des propositions: "Si m est une classe de propositions, la proposition < tout m est vrai > peut être ou ne pas être elle-même un m . Mais il y a une relation biunivoque de cette proposition à m ; si n est différent de m , < tout n est vrai > n'est pas la même proposition que < tout m est vrai >. Considérons maintenant la classe totale des propositions de la forme < tout m est vrai > et ayant la propriété de ne pas être élément de leur m respectif. Appelons w cette classe et soit p la proposition < tout est vrai >. Si p est un m , il doit posséder la propriété définissante de w et donc est un w . Ainsi la contradiction apparaît inévitable"⁶.

On a, d'autre part, de la forme (A) dérivé le Menteur⁷.

$$(A_M) \{ [(x)(x \varepsilon F \supset \langle x \varepsilon F \rangle \not\equiv F) \supset [q \not\equiv F]] \supset \\ \{ [(x)(\langle x \varepsilon F \rangle \not\equiv F \equiv x \varepsilon F) \supset (q \not\equiv F \cdot q \varepsilon F)] \}$$

L'antinomie (A_p) a été obtenue à partir de la matrice générale (A) par les assignations:

$$(1) \quad \begin{cases} \phi_x = (\exists q_x) (x \leftrightarrow q_x \cdot x \not\equiv q_x) \\ f_k^s = p_k \end{cases}$$

D'autre part, la correspondance biunivoque entre les classes de propositions m et les propositions de la forme < tout m est vrai > est garantie par la propriété:

$$p_{m_q} = q \quad \text{et} \quad q_{p_m} = m.$$

Sans plus tenir compte alors de la signification prêtée par Russell à la forme restituée (A_p), les assignations (1) donnent l'écriture complète:

$$(A_p)' (u) \{ [(x)(x \varepsilon u \supset (\exists q_x)(x \leftrightarrow q_x \cdot x \not\equiv q_x)) \supset (p_u \leftrightarrow u \cdot p_u \not\equiv u)] \supset \\ \{ [(\exists w)(y) [(\exists q_y)(y \leftrightarrow q_y \cdot y \not\equiv q_y) \equiv y \varepsilon w]] \supset \\ [(\exists w)(p_w \leftrightarrow w \cdot p_w \not\equiv w \cdot p_w \varepsilon w)] \}$$

Interprétons les assignations (1) de la façon suivante:

" x a la propriété ϕ si et seulement si il existe un prédicat q dont la définition est associée biunivoquement (par exemple par un arrangement lexicographique, ou, dans une syntaxe arithmétisée, par le nombre correspondant au prédicat) à x et qui ne convient pas à x "; " $f^s u$ est le nombre p_u , associé au prédicat définissant de u ". Dans cette interprétation, les q_x sont des signes de prédicats, les x et les p_u des signes numériques (des numéros), les u des ensembles de nombres.

On obtient alors la formulation du paradoxe de Richard: "Si, quel que soit x , si, lorsque x est un u , il existe un prédicat q dont la définition soit associée biunivoquement à x et qui ne convienne pas à x , le nombre associé à u n'est pas un u , alors, s'il existe une classe w de tous les y tels qu'il existe un prédicat q dont la définition soit associée biunivoquement à y et

6. Russell, *Principles of Mathematics*, Appendice B, p. 527.

7. Vuillemin, *op. cit.*, p. 85.

qui ne convienne pas à y , le nombre associé à cette classe w lui appartient et en même temps ne lui appartient pas”.

Or on peut faire subir à cette matrice (A_p) les mêmes simplifications que tout à l’heure.

Carnap résume ainsi l’antinomie de Richard: “L’univocité du numérotage (des définitions des prédicats) est supposée⁸ :

$$(\text{num}(F) = \text{num}(G)) \supset (x) (F(x) \equiv G(x)). \quad (1)$$

Avec l’aide de ‘num’, ‘Ri’ (“Richardien”) peut maintenant être défini:

$$\text{Ri}(x) \equiv (F) [(\text{num}(F) = x) \supset \sim F(x)]. \quad (2)$$

Puisque ‘Ri’ est un prédicat numérique à une place, il a un certain numéro particulier désigné par ‘num(Ri)’. Nous supposons d’abord que le numéro de ‘Ri’ est lui-même Richardien: $\text{Ri}[\text{num}(\text{Ri})]$. Alors, si nous substituons en (2) ‘num(Ri)’ à la place de ‘x’, et ‘Ri’ à la place de ‘F’, ‘[Ri num(Ri)]’ suit aisément. Puisque notre supposition conduit à son contraire, il s’ensuit qu’elle est réfutée; et donc il est prouvé que

$$\sim \text{Ri}[\text{num}(\text{Ri})]. \quad (3)''$$

Cette partie du raisonnement de Carnap correspond au premier membre de l’implication de (A_p) , soit:

$$(A_{p2})' [(u) [(x)(x \varepsilon u \supset ((\exists q_x)(x \leftrightarrow q_x \cdot x \not\# q_x)) \supset p_u \not\# u)],$$

si l’on admet les correspondances: $x \varepsilon u \Rightarrow \text{Ri}(x)$, $(\exists q_x) x \leftrightarrow q_x \Rightarrow (F) (\text{num}(F) = x)$, $x \not\# q_x \Rightarrow \sim F(x)$. La généralisation (F) , chez Carnap, assure automatiquement que Ri, qui est un F , possède un numéro associé; cette conséquence résulte en (A_p) du jeu combiné des conditions $(\phi(f'u) \cdot f'u \not\# u)$ avec les assignations (1).

La seconde partie du raisonnement de Carnap correspond au conséquent de (A_p) , soit:

$$(A_{p1})' \{(\exists w)(y) [(\exists q_y)(y \leftrightarrow q_y \cdot y \not\# q_y) \equiv y \varepsilon w]\} \supset [(\exists w)(p_w \leftrightarrow w \cdot p_w \not\# w \cdot p_w \varepsilon w)].$$

En effet: “Par (1):

$$(\text{num}(F) = \text{num}(\text{Ri})) \supset (\sim F[\text{num}(\text{Ri})] \equiv \sim \text{Ri}[\text{num}(\text{Ri})]). \quad (4)$$

Par (3), (4):

$$(\text{num}(F) = \text{num}(\text{Ri})) \supset \sim F[\text{num}(\text{Ri})]. \quad (5)$$

Par [2]:

$$(F)[(\text{num}(F) = \text{num}(\text{Ri})) \supset \sim F[\text{num}(\text{Ri})]] \supset \text{Ri}[\text{num}(\text{Ri})]. \quad (6)$$

Par (5), (6):

$$\text{Ri}[\text{num}(\text{Ri})] \quad (7)$$

Les sentences démontrées (3) et (7) se contredisent l’une l’autre”⁹.

8. ‘num’ signifie ‘le numéro de’

9. Carnap, *op. cit.*, pp. 219-220.

On peut donc simplifier $(A_p)'$ en $(A_{p1})'$. L'adjonction de $(A_{p2})'$ a pour seul effet de supprimer la forme dilemmatique du raisonnement, en éliminant d'emblée l'hypothèse selon laquelle le nombre qui correspond à la propriété "Richardien" par le numérotage des définitions pouvait être lui-même richardien.

Naturellement, on pourrait apporter la même simplification d'écriture à (A_p) et à (A_M) . Pour (A_{p1}) et $(A_{p1})'$, on ferait les mêmes remarques que tout à l'heure sur la forme fermée de ces propositions comparée à la forme couverte préférée par Carnap et sur la différence entre quantification existentielle et constante. Il en irait encore de même pour (A_{M1}) :

$$[(x) (\langle x \varepsilon F \rangle \not\equiv x \varepsilon F) \supset (q \not\equiv F \cdot q \varepsilon F)]$$

en ce qui concerne la liaison de x .

III. De ces remarques, il ne faudrait pas conclure que les antinomies sémantiques - à la différence des antinomies logiques - n'admettent les simplifications qu'à partir de la forme générale (A) et non à partir de la forme réduite (B).

L'antonomie de Grelling fournirait un contre-exemple à cette supposition.

"Définition: dans un langage qui contient sa propre syntaxe, un prédicat syntactique (par exemple, un adjectif) est dit *hétérologique* si la sentence qui attribue la propriété exprimée par le prédicat au prédicat lui-même est fautive. Si, par exemple, 'Q' est un prédicat syntactique, alors "Het ('Q') $\equiv \sim Q('Q')$ " est vraie." Exemple: l'adjectif 'monosyllabique' est hétérologique, parce que 'monosyllabique' n'est pas monosyllabique mais hexa-syllabique. Maintenant, si au lieu du prédicat 'Q', nous prenons le prédicat 'Het' lui-même, qui vient d'être défini, nous obtenons, à partir de la définition établie, la sentence contradictoire "Het (Het') $\equiv \sim \text{Het}(\text{Het}')$ "¹⁰. On dériverait directement cette antonomie de la forme simplifiée (B₁).

Ainsi, toutes les antinomies, soit logiques soit sémantiques, peuvent se déduire des matrices simplifiées et devenues dilemmatiques:

$$(A_1) [(\exists w)(y)(\phi y \equiv y \varepsilon w)] \supset [(\exists w)(\phi f'w \cdot \sim \phi f'w)]$$

$$(B_1) [(\exists w)(y)(y \not\equiv y \varepsilon w)] \supset [(\exists w)(w \not\equiv w \cdot w \varepsilon w)].$$

On peut dans ces matrices 1^o substituer aux propositions énonçant l'appartenance d'un élément à une classe des propositions prédicatives¹¹, 2^o éliminer les quantificateurs universels, 3^o remplacer le quantificateur existentiel $(\exists w) \dots \equiv y \varepsilon w$ par une constante.

§2. Le théorème de Gödel et la matrice de l'antonomie

Pour dériver le théorème de Gödel à partir de la matrice de

10. Carnap, *op. cit.*, § 60 a, pp. 211-212.

11. Au sens: énonçant qu'une propriété convient à un élément.

l'antinomie, nous nous référerons à la présentation abrégée que donne Gödel de sa démonstration.¹²

Soit le formalisme L des *Principia Mathematica*, supposé cohérent, arithmétisé par Gödel. Si q est une proposition de L , l'arithmétisation permet d'exprimer sous la forme d'un énoncé de l'Arithmétique récursive la proposition: " q est dérivable en L ". D'autre part, l'ensemble des propositions de L contenant une variable libre individuelle susceptible d'être remplacée par un chiffre est dénombrable; on peut donc l'ordonner par une relation R de l'Arithmétique récursive, représentable en L par le prédicat R^* . Chaque proposition $q(x)$ de L reçoit ainsi un numéro. Nous pouvons écrire l'ensemble de ces propositions sous la forme de la suite:

$$q_1(x), q_2(x), \dots, q_n(x), \dots$$

A chacune de ces propositions ouvertes correspond une suite infinie de propositions fermées, obtenues en substituant un chiffre à la variable:

$$\begin{array}{l} q_1(x) \leftrightarrow q_1(1), q_1(2), q_1(3), \dots, q_1(n), \dots \\ q_2(x) \leftrightarrow q_2(1), q_2(2), q_2(3), \dots, q_2(n), \dots \\ q_3(x) \leftrightarrow q_3(1), q_3(2), q_3(3), \dots, q_3(n), \dots \\ \cdot \\ \cdot \\ \cdot \\ q_n(x) \leftrightarrow q_n(1), q_n(2), q_n(3), \dots, q_n(n), \dots \end{array}$$

Considérons la suite diagonale:

$$q_1(1), q_2(2), q_3(3), \dots, q_n(n), \dots$$

Toutes les propositions fermées étant ou dérivables ou non-dérivables dans L , une bipartition de l'ensemble diagonal s'ensuit. Soit alors le sous-ensemble de la suite constitué par l'ensemble des propositions non-dérivables dans L . Nommons u la classe des numéros associés par l'ordre R à ces propositions non dérivables. Si nous appelons nombre gödelien d'une proposition le numéro qui lui est associé par R et si nous désignons par Dém $q_n(x)$ l'énoncé métathéorique: "La proposition de numéro gödelien n est dérivable en L ", on obtient donc:

$$(1) \quad n \in u \equiv \sim \text{Dém } q_n(n).$$

En d'autres termes, le nombre n appartient à la classe u si et seulement si la proposition obtenue en remplaçant la variable x de $q_n(x)$ par le symbole représentant le nombre n , n'est pas dérivable dans L .

12. Kurt Gödel, *Ueber formal unentscheidbare Sätze der Principia Mathematica und Verwandter Systeme*, Monatshefte für Mathematik und Physik, vol. 38, 1931, pp. 124-176; Jean Ladrière, *Les Limitations internes des formalismes*, Nauwelaerts. Gauthier-Villars, Louvain, Paris, 1957, §72, pp. 102-104 et §237, pp. 418-420.

Or la propriété d'appartenir à la classe u est représentable en L par une expression r - jouant le rôle de prédicat - . L'expression $r(x)$ désigne la proposition ouverte: " x appartient à la classe u ". Cette proposition est numérotée par R^* et on peut donc l'écrire, si ce numéro est s , $r_s(x)$. Nous pouvons alors former, par le même procédé qu'auparavant, la proposition $r_s(s)$, qui appartient à L et qui signifie:

"Le nombre entier s appartient à la classe u ",

c'est-à-dire, en vertu de (1):

"La proposition obtenue, en remplaçant la variable x de $r_s(x)$ par le symbole qui représente s , n'est pas dérivable dans L ",

c'est-à-dire:

"La proposition $r_s(s)$ n'est pas dérivable dans L ".

La Proposition $r_s(s)$ affirme donc sa propre indériverabilité dans L .

En effet, supposons que $r_s(s)$ soit dérivable. Alors l'entier s a la propriété énoncée par la proposition $r_s(x)$ et donc fait partie de u . Mais alors: \sim Dém $r_s(s)$. - Supposons que $r_s(s)$ soit réfutable. Alors s n'appartient pas à u et Dém $r_s(s)$. Donc si $r_s(s)$ est soit dérivable, soit réfutable, L est incohérent. Mais on a supposé que L était cohérent. Donc $r_s(s)$ est indériverable.

L'analogie avec la démonstration utilisée dans le théorème de Cantor et avec la numérotation utilisée dans l'antinomie de Richard permet de reconstruire aisément le théorème de Gödel à partir de l'antinomie de forme générale (A), par les assignations:

$$(2) \quad \begin{cases} \phi x = (\exists q_x(x)) (x \leftrightarrow q_x(x) \cdot \sim \text{Dém } q_x(x)) . \\ f'_u = (\exists m) R^*(x \varepsilon u, m) \cdot m \leftrightarrow q_m(x) . \end{cases}$$

On interprétera les variables m, n, o, p comme des variables numériques (substituables par les chiffres, c'est-à-dire les constantes de L), les variables x comme des variables d'individus et q, r, s, t comme des propositions. Alors la matrice du théorème a pour expression:

$$(A_G) \quad \{ \{ u \} [(m) m \varepsilon u \supset (\exists q_m(m)) (m \leftrightarrow q_m(m) \cdot \sim \text{Dém } q_m(m))] \supset \\ [(\exists n) R^*(x \varepsilon u, n) \cdot n \leftrightarrow r_n(x) \cdot \sim \text{Dém } r_n(n)] \} \supset \\ \{ [(\exists w)(o) [(\exists s_o(o)) (o \leftrightarrow s_o(o) \cdot \sim \text{Dém } s_o(o)) \equiv o \varepsilon w] \supset \\ [(\exists w) (\exists p) R^*(x \varepsilon w, p) \cdot p \leftrightarrow t_p(x) \cdot \sim \text{Dém } t_p(p) \cdot \sim \text{Dém } t_p(p)] \} \} .$$

L'expression peut être simplifiée et le raisonnement prend alors une forme dilemmatique:

$$(A_{G1}) \quad \{ (\exists w)(o) [(\exists s_o(o)) (o \leftrightarrow s_o(o) \cdot \sim \text{Dém } s_o(o)) \equiv o \varepsilon w] \supset \\ [(\exists w) (\exists p) R^*(x \varepsilon w, p) \cdot p \leftrightarrow t_p(x) \cdot \sim \text{Dém } t_p(p) \cdot \sim \text{Dém } t_p(p)] \}$$

§3. Distinction des antinomies logiques et sémantiques: Propriétés et prédicats.

Si l'on excepte le cas de l'antinomie de Grelling¹³, la différence entre

13. Voir plus haut, p. 6.

les antinomies logiques et les antinomies sémantiques n'a pas paru dans nos notations. En conséquence, puisque ce sont les secondes qui "inspirent" le raisonnement de Gödel,¹⁴ elles n'ont pas non plus paru dans les expressions (A_G) et (A_{G1}) .

Or, en quoi consiste cette différence?

Revenons à l'antinomie logique de l'Imprédictible et comparons-la à l'antinomie sémantique de l'Hétérologique. Bien que, dans les deux cas, j'obtienne l'antinomie en appliquant Imprédictible et Hétérologique à eux-mêmes, cette application se fait en deux sens différents. Lorsque je dis que l'Imprédictible est imprédictible, j'attribue la propriété "Imprédictible" à la *propriété* "Imprédictible". Mais lorsque je dis qu'Hétérologique est hétérologique, j'attribue la propriété "Hétérologique" non pas à la propriété "Hétérologique", mais au *prédicat* "Hétérologique", le prédicat étant ici entendu comme la désignation de (le symbole désignant) la propriété.

L'usage correct de l'écriture impose de distinguer ces deux modes d'attributions. Dans le premier cas, où le sujet de l'attribution est une *suppositio formalis*, le prédicat qui la désigne est utilisé - préfixé de l'article - et écrit ordinairement:

"Le concept est un concept":

Dans le second cas, où le sujet de l'attribution est une *suppositio materialis*, le prédicat qui la désigne est utilisé entre guillemets:

" "Monosyllabique" n'est pas monosyllabique".

Cette différence peut encore s'exprimer dans les genres des substitutions qui conduisent aux deux sortes d'antinomies. Lorsque, dans la matrice (B_3) nous obtenons la contradiction exprimée dans le membre droit de l'implication en substituant "Impr" à toutes les occurrences de "F" dans le membre gauche, la triple substitution qui a lieu est homogène. Au contraire, dans la matrice du paradoxe de Grelling:

$$[\text{Het} ('Q') \equiv \sim Q('Q')] \supset [\text{Het} ('Het') \equiv \sim \text{Het} ('Het')],$$

nous obtenons la contradiction exprimée dans le membre droit de l'implication en substituant "Het" à toutes les occurrences de "F" dans le membre gauche, mais deux de ces substitutions sont formelles, tandis que la troisième est matérielle.

Souvent, les auteurs mêmes qui insistent sur la nécessité d'un langage rigoureux violent la prescription, qui ordonne de respecter cette restriction.

14. "Les sentences syntactiques peuvent parfois parler sur elles-mêmes et la question se pose de savoir si cette réflexivité peut éventuellement conduire à des contradictions. Cette question est importante, parce qu'elle se rapporte non pas à des calculs d'un type spécial de construction, mais à tous les systèmes, quels qu'ils soient, qui contiennent l'Arithmétique . . . Le problème des *antinomies syntactiques* . . . réapparaît visiblement quand il est question d'un langage S, dans lequel la syntaxe de S lui-même peut être formulée, et dans le cas de tout langage qui contient l'Arithmétique" (Carnap, *op. cit.*, §60 a, pp. 211-213).

Dans la matrice (A_P) et (A_M) , on évitait ces difficultés en recourant au langage des classes. Mais dans la matrice $(A_P)'$, c'est bien un *prédicat* " q_x " et non une *propriété* " q_x " qui intervient dans la numérotation, même si c'est la *propriété* " q_x " qui est prédiquée du numéro qui lui correspond. Une écriture correcte de l'antinomie exigerait donc, par exemple, pour $(A_{P1})'$, les signes:

$$\{(\exists w)(y) [(\exists q_y)(y \leftrightarrow "q_y" \cdot y \not\vdash q_y) \equiv y \varepsilon w]\} \supset [(\exists w)(p_w \leftrightarrow w \cdot p_w \not\vdash w \cdot p_w \varepsilon w)].$$

L'avantage qu'il y a à remplacer le quantificateur existentiel par une constante, comme fait Carnap, apparaît ici en ce que ce procédé supprime la question de savoir si la liaison existentielle porte sur une variable symbole de propriété ou sur une variable symbole de prédicat. De même, si l'on interprète " F " comme un prédicat et non comme la classe du faux, on devra écrire la matrice simplifiée du menteur (A_{M1}) de la façon suivante:

$$[(x) ("x' \varepsilon F' \not\vdash F' \equiv "x' \varepsilon F") \supset ("q' \not\vdash F.' "q' \varepsilon F")]$$

De même, dans la matrice de Gödel (A_{G1}) , il est évident que la relation du numérotage $R^*(x \varepsilon w, p)$ doit s'écrire: $R^*(x \varepsilon w', p)$.

Ces distinctions ont une double portée. Elles permettent de légitimer le raisonnement de Gödel et de démontrer d'autres théorèmes de limitation. Elles font apparaître que les antinomies sémantiques sont liées non pas à la propriété logique de réflexivité, mais à la capacité "universelle" d'expression d'une langue.

§4. *Légitimité de la démonstration de Gödel; autres théorèmes de limitation.*

On a pu remarquer¹⁵ que, dans l'énoncé de la matrice (A_G)

$$o \varepsilon w \equiv \sim \text{Dem } s_o(o),$$

qui établit, en se fondant sur la numérotation des propositions de L une certaine relation entre la classe w et la dérivabilité de la proposition $s_o(o)$, "l'expression $s_o(o)$ est utilisée selon l'usage autonome. Elle ne doit donc pas être considérée, dans ce contexte, comme une expression de L , mais comme un nom appartenant à la métathéorie de L et servant à désigner une certaine expression de P . On pourrait éviter toute ambiguïté en utilisant une notation spéciale ou en recourant à une correspondance de Gödel".

15. Ladrrière, *op. cit.*, pp. 142-143, à propos de l'essai de Ch. Perelman pour réduire la démonstration de Gödel à une nouvelle forme d'antinomie. De même, sur le même sujet: Kurt Grelling, *Bemerkungen zu einer Abhandlung von Ch. Perelman*, *Theoria*, III, 1939, pp. 297-306 (p. 300: "Mais $x \varepsilon E$ (dans la matrice (AG) : $m \varepsilon u$ ou $x \varepsilon w$, etc.) n'est pas pour autant un signe de classe (c'est-à-dire une formule de P - langue formelle sur laquelle raisonne Gödel -, qui exprime d'une façon quelconque une propriété des nombres naturels); car un tel signe devrait pouvoir être exprimé dans le système P ; $x \varepsilon E$ n'est qu'une description méta-mathématique d'une propriété numérique".

Dans L , la proposition $s_0(o)$ est associée à un numéro gödelien et la proposition selon laquelle elle n'est pas dérivable est exprimée par l'image arithmétique de la proposition métathéorique $s_0(o)$.

On devra donc écrire la matrice (A_{G_1}) de la façon suivante:

$$\{(\exists w)(o) [(\exists s_0(o)(o \leftrightarrow 's_0(o)' \cdot \sim \text{Dém } 's_0(o)')] \equiv o \varepsilon w\} \supset \\ \{(\exists w)(\exists p) [R^*(x \varepsilon w', p) \cdot p \leftrightarrow 't_p(x)' \cdot \sim \text{Dém } 't_p(p)' \cdot \sim \text{Dém } '\sim t_p(p)']\}.$$

Ainsi, de même qu'à la matrice (A_p) du théorème de Cantor correspond la matrice (A) des antinomies logiques, de même, à la matrice (A_G) du théorème de Gödel correspond la matrice ($A_{p'}$) de l'antimie de Richard. On a ailleurs¹⁶ montré que, pour passer du raisonnement cantorien aux antinomies logiques, il suffit de considérer l'ensemble diagonal L , obtenu en prenant les éléments de l'ensemble M qui ne figurent pas dans le sous-ensemble P_m de $\mathcal{P}(M)$ associé biunivoquement à M , et de le forcer à être élément de M ; autrement dit, il suffit de prendre pour M la classe universelle V .

Examinons à présent comment on passe de même du théorème de Gödel à l'antimie de Richard. Deux remarques complémentaires de Carnap et de Church éclairent cette question.

Si, dans l'antimie du menteur, on remplace "faux" par "non-démonstrable", nous obtenons une sentence S_1 qui énonce d'elle-même qu'elle n'est pas démontrable dans le langage choisi, soit S . Cette sentence est l'analogue de la proposition gödelienne. Ici, aucune contradiction ne se produit. Si S_1 est vraie (analytique), alors S_1 n'est pas fautive (contradictoire), mais seulement non démontrable en S . C'est, en fait, le cas. Les propriétés "analytiques" et "non-démonstrable" ne sont pas incompatibles.

Replaçons, à présent, "faux" par "réfutable" dans la sentence du menteur. Supposons qu'une sentence, S_2 soit construite en S , qui affirme que S_2 est elle-même réfutable (en S). S_2 est alors un analogue à l'assertion du menteur. Observons alors si la contradiction se produit selon le mode ordinaire. Supposons d'abord que S_2 soit en fait réfutable. Alors S_2 sera vraie et donc analytique. D'autre part, toutefois, toute sentence réfutable est contradictoire, et donc non analytique. Donc la supposition est fautive et S_2 est non-réfutable. De ceci ne résulte aucune contradiction. S_2 est en fait non réfutable; puisque S_2 signifie le contraire de ceci, S_2 est fautive et est donc contradictoire - Mais les propriétés "non-réfutable" et "contradictoire" sont tout à fait compatibles l'une avec l'autre.¹⁷

Soit, d'autre part, l'énoncé de Findlay, qui construit dans le langage quotidien un énoncé circulaire correspondant à la proposition indécidable de Gödel: "Il est impossible de démontrer l'énoncé de fait que l'on obtient en remplaçant, dans la formule énonciative: <Il est impossible de démontrer l'énoncé de fait que l'on obtient en remplaçant, dans la formule énonciative A , la variable qu'elle contient par le nom de cette formule

16. Vuillemin, *op. cit.*, §9, p. 87.

17. Carnap, *op. cit.*, §60c, pp. 217-218.

énonciative>, la variable qu'elle contient par le nom de cette formule énonciative".¹⁸ Si, dans cet énoncé, on substitue le prédicat *non-vrai* à la phrase *Il est impossible de démontrer* (au prédicat *non-démontrable*), on obtient un énoncé qui ressemble à l'antinomie d'Epiménide.¹⁹ Or, on obtient directement, à partir de la matrice (A_{G1}), la matrice simplifiée de cette antinomie, en y substituant " \sim Vrai $q_m(m)$ " à " \sim Dém $q_m(m)$ ":

$$(A_E) \{ (\exists w)(o) [(\exists s_0(o)) (o \leftrightarrow "s_0(o)" \cdot \text{Faux } "s_0(o)")] \equiv o \varepsilon w \} \supset \\ [(\exists w)(\exists p) R^*("x \varepsilon w", p) \cdot p \leftrightarrow "t_p(x)" \cdot \text{Faux } "t_p(p)" \cdot \text{Faux } "\sim t_p(p)"],$$

en posant:

$$" \sim \text{Vrai } ("q") = \text{Faux } ("q") " .$$

On notera que (A_E) est la matrice de l'Epiménide. Elle diffère par l'existence des quantificateur de la matrice du menteur (A_{M1}). Elle énonce que si un nombre appartient à w si et seulement si il est associé à une proposition obtenue en remplaçant la variable libre par le symbole représentant le nombre associé et disant d'elle-même qu'elle est fausse, alors il existe un nombre p associé à la proposition t disant que $x \varepsilon p$, et la proposition obtenue de t en substituant à la variable libre le symbole représentant p dit d'elle-même qu'elle est fausse et qu'elle ne l'est pas.²⁰

Allons plus loin.

Le théorème de Cantor a pour fin de rejeter, en vertu d'un raisonnement par l'absurde l'hypothèse selon laquelle l'ensemble diagonal, L , qu'on construit appartiendrait à l'ensemble de départ, M . Or, dans la matrice (A_p)' de l'antinomie de Richard, la contradiction dépend d'une double hypothèse faite sur $y \varepsilon w$. D'une part le nombre associé par la numérotation ne doit pas vérifier la propriété définie. De l'autre, la numérotation doit être effectuable de façon univoque dans la langue formelle choisie, S . Ainsi le paradoxe de Richard peut être transformé en théorème, en supposant que la langue S est cohérente: "Si S est consistante ou du moins non contradictoire, alors il n'est pas possible de construire en S ni un cadre expressif ni une expression fonctorielle, en vertu desquels on pourrait construire en S une énumération univoque des prédicats d'arguments numériques à une place. Bien que l'ensemble des prédicats d'arguments numériques à une place, qui sont définissables en S , soit un ensemble dénombrable, . . . , leur énumération ne peut pas être effectuée par les moyens dont on dispose en S même".²¹

18. J. N. Findlay, *Gödelian-sentences: a non numerical approach*, Mind, n. 5, vol. 51, 1942, pp. 259-265; Ladrière, *op. cit.*, p. 467.

19. A Church, Compte-rendu de l'article de Findlay, *Journal of Symbolic Logic*, vol. 7, 1942, pp. 129-130; Ladrière, *op. cit.*, p. 117.

20. L'Epiménide revient donc à dire non pas qu'un Crétois, Epiménide, dit que tous les Crétois sont menteurs, forme qui ne conduit pas à une antinomie (S. C. Kleene, *Introduction to Metamathematics*, p. 39), mais que tous les Crétois disent que tous les Crétois sont menteurs, l'ensemble des Crétois étant supposé non-vide.

21. Carnap, *op. cit.*, §60c, Théorème 60c. 3, p. 220.

De l'antinomie de Grelling, on peut tirer des conclusions analogues.²² D'ailleurs, pour obtenir cette antinomie à partir de (A_{Gr1}), il suffit de substituer, dans cette matrice, le prédicat 'An' (= 'Analytique en L') au prédicat 'Dém' (= 'Démonstrable en L'). On obtient alors la matrice:

$$(A_{Gr1}) \{ (\exists w)(o) [(\exists s_o(o))(o \rightarrow 's_o(o)' \cdot \sim An 's_o(o)'] \equiv o \varepsilon w \} \supset \\ \{ (\exists w)(\exists p) [R^*('x \varepsilon w', p) \cdot p \leftrightarrow 't_p(p)' \cdot \sim An 't_p(p)' \cdot \sim An '\sim t_p(p)'] \}.$$

Il suffit alors de définir:

$$Het(x) =_{Df} (\exists q_x(x)) (x \leftrightarrow 'q_x(x)' \cdot \sim An 'q_x(x)'),$$

ce qui revient, dans l'assignation (2) à remplacer ϕx par $Het(x)$, pour retrouver, dans la forme, l'antinomie de Grelling. Mais, c'est constater ainsi que si "analytique en S" est définissable en S, S est contradictoire. De (A_{Gr1}), on tirera donc le théorème suivant: "Si S est consistante ou du-moins non contradictoire, alors "analytique (en S)" est indéfinissable en S".²³

Le célèbre Corollaire de Gödel est étroitement lié à ces théorèmes, lorsqu'on le formule sémantiquement.²⁴ En vertu de ce Corollaire, si S est consistante ou du moins non contradictoire, alors aucune preuve de la non contradiction de S ne peut être formulée dans une syntaxe qui n'utilise que les moyens d'expression dont on dispose en S.

§5. *Classification des théorèmes de limitation: Dérivation et conséquence*

Avant d'examiner la raison d'être des faits qu'on a décrits et d'expliquer pourquoi et comment on peut, des matrices d'antinomies, déduire des matrices correspondantes de théorèmes de limitation, il convient de classer ces faits et ces théorèmes.

On observera d'abord que les expressions "faux", "vrai", "définissable", "indéfinissable" appartiennent à la Sémantique. Les deux concepts fondamentaux de cette science sont, en effet celui de dénotation et celui de valeur²⁵; le premier circonscrit la théorie de la Définition, le second celui de la Vérité. Ainsi les trois antinomies du menteur (A_M), de l'Epiménide (A_E) et de Richard (A_p) peuvent être dites antinomies sémantiques.

En second lieu, Tarski et Carnap ont résolu, pour différents langages formels, le problème de trouver une relation syntactique qui coïncide en extension avec la relation sémantique: "satisfaire une forme propositionnelle".²⁶ Par exemple, à la notion sémantique de "Vrai", on peut faire correspondre la notion syntactique d'"Analytique" ou de "Valide",²⁷

22. Carnap, *op. cit.*, §60 c, p. 218.

23. Carnap, *op. cit.*, §60 c, Théorème 60 c. 1, p. 219.

24. Carnap, *op. cit.*, §60 c, p. 219.

25. A. Church, *Introduction to Mathematical Logic*, §09, note 143, p. 65.

26. A. Church, *Ibid.*, note 142, p. 65.

27. Ces deux concepts sont considérés comme identiques, lorsque les langues formelles sur lesquelles on raisonne sont purement logiques. Elles ne se distinguent que dans le cas langages "physiques" (Carnap, *op. cit.*, §48, p. 173 et §51, pp. 180-182), qui ne nous intéressent pas ici.

chaque fois que le faux est tel en vertu des seules lois de la Logique et non en vertu d'un fait; de même, pour les mêmes conditions, on fera correspondre la notion syntactique de "Contradictoire" ou de "Contravalide" à la notion sémantique de "Logiquement Faux". On notera que ces concepts syntactiques ne sont pas "complets", en ce sens que, dans le cas général, nous ne possédons pas de méthode de décision pour déterminer si une proposition donnée est effectivement analytique ou contradictoire. Elle l'est en soi; je ne sais pas nécessairement si elle l'est. Nous appellerons avec Carnap²⁸ c-termes (termes de la méthode de la conséquence) les termes syntactiques ainsi sélectionnés. L'antinomie de Grelling (A_{Gr}) appartient à ce groupe.

Enfin, on appellera d-termes (termes de la méthode de la dérivation) les termes syntactiques tels que: "Dérivable", "Démonstrable", "Réfutuable", "Résoluble", "Irrésoluble", qui exigent une méthode de décision. Le théorème de Gödel (A_G) appartient à ce groupe.

Quand on peut faire correspondre un concept syntactique à un concept sémantique, on obtient, à partir d'une antinomie ou d'un théorème sémantique une antinomie ou un théorème syntactique. Par exemple, on obtient une version syntactique du menteur et de l'Epiménide, en remplaçant, dans les matrices de ces antinomies, les mots "Faux" par "Contradictoire".

Or à toute antinomie sémantique ou à toute antinomie syntactique en c-termes, à condition qu'une règle d'énumération R^* existe dans le langage formel S pour les expressions convenablement choisies de S , correspond un théorème sémantique de limitation. A l'Epiménide (A_E), correspond le théorème de la vérité de Tarski ou le Corollaire de Gödel exprimé sous forme sémantique à la condition, dans ce dernier cas, qu'on substitue à "Faux" le prédicat syntactique "Contradictoire". En vertu du théorème de la vérité, si un formalisme L est tel qu'on y puisse formaliser l'Arithmétique ordinaire, les règles de la Logique et la Théorie des ensembles et qu'on y considère la classe K des propositions vraies de L , contenant les axiomes de L et fermée pour la relation de conséquence, cohérente et complète, alors la théorie du prédicat "Vrai" n'est pas formalisable dans le système L lui-même. A l'antinomie de Grelling correspond le théorème affirmant l'impossibilité de définir dans le formalisme considéré le prédicat "Analytique", ce théorème n'étant que le doublet du théorème de Tarski. A l'antinomie de Richard correspond, de même, le théorème sur l'impossibilité d'énumérer les prédicats de L dans L .

Tous ces théorèmes expriment des limitations des formalismes eu égard à la possibilité d'exprimer certains concepts fondamentaux de la syntaxe ou de la sémantique de ces formalismes dans ces formalismes eux-mêmes. Ils diffèrent des théorèmes proprement syntactiques, tels le théorème de Gödel et son corollaire, qui expriment des limitations des

28. Carnap, *op. cit.*, §34 a, p. 101.

formalismes eu égard à la possibilité de trouver une méthode de décision pour toutes les sentences exprimables dans ces formalismes.

Ces deux sortes de théorèmes font donc peser deux sortes de limitations sur les systèmes arithmétiques.²⁹ Les unes établissent que certains termes arithmétiques sont indéfinissables - naturellement abstraction faite des indéfinissables reconnus par le système, en tant qu'ils en constituent les notions primitives. Les autres établissent que certaines sentences arithmétiques sont indécidables, c'est-à-dire qu'on ne peut les prouver à partir des axiomes, - eux-mêmes assertés quoiqu'indémontrables *ex hypothesi*, - bien qu'on puisse les exprimer en vertu des règles de formation.

§6. *Réflexivité, antinomie et capacité d'expression d'une langue*

On a cru souvent que les antinomies étaient liées au caractère "réflexif" de certaines expressions. Mais, d'une part on peut construire des antinomies qui correspondent à celles du menteur, sans utiliser la relation directe de réflexivité qui rapporte une proposition à elle-même³⁰. De l'autre, on peut construire, comme l'ont fait Gödel et Tarski, des propositions "réflexives" qui ne sont pas antinomiques.

"Ce n'est pas, comme on l'a cru si souvent, la réflexivité qui constitue l'erreur de laquelle la contradiction dépend; l'erreur se trouve plutôt dans l'usage sans restriction des termes 'vrai' et 'faux'".³¹ C'était déjà la généralité du mot *tous* qui instaurait les antinomies logiques. C'est la généralité des concepts sémantiques (vrai/faux, définissable/indéfinissable, énumérable/non énumérable), ainsi que des termes syntactiques qui peuvent leur correspondre, sans égard pour le langage de référence auquel on rapporte ces concepts, qui est, de même, responsable des antinomies sémantiques.

Si l'on a pu faire se correspondre les matrices (A_G) et (A_E) et obtenir dans un cas ou le théorème syntactique de Gödel ou le théorème sémantique de Tarski et dans l'autre l'Epiménide, c'est que, dans le premier cas, on opérait sur des concepts syntactiques effectifs et limités, tels que "Démontrable" et "Non-Démontrable", tandis que, dans l'Epiménide, interviennent des concepts sémantiques (ou syntactiques) non effectifs et illimités, tels que le "Vrai" et le "Faux" (l'"Analytique" et le "Contradictoire") et que, dans le théorème de Tarski, on se fondait sur la contradiction produite par l'usage de tels concepts pour en interdire la définissabilité dans un langage formel supposé cohérent.

Comme le note Carnap, "l'impossibilité de reconstruire l'antinomie du

29. Carnap, *op. cit.*, §60d, Théorème 2, p. 221.

30. Carnap, *op. cit.*, §60b, pp. 215-216; de même, Lewis et Langford, *Symbolic Logic*. On notera une situation semblable, au point de vue mathématique, dans le cas des ensembles extraordinaires (D. Mirimanoff, *Les antinomies de Russell et de Burali-Forti et le problème fondamental de la théorie des ensembles*, L'Enseignement mathématique, tome XIX, p. 37-52, 1917).

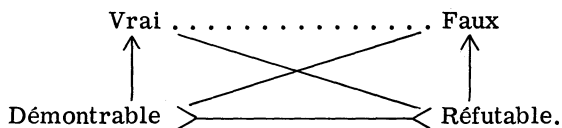
31. Carnap, *op. cit.*, §60b, p. 215.

Menteur à l'aide des termes <non-démontrable> ou <réfutable> est due au fait que toutes les sentences analytiques ne sont pas démontrables et que toutes les sentences contradictoires, de même, ne sont pas toutes réfutables".³² Et si, inversement, l'on peut reconstruire l'Epiménide à partir du théorème de Gödel, ce phénomène, note Church, tient, de même, "au fait qu'un système peut très bien ne pas contenir une formalisation de sa sémantique et contenir cependant une formalisation d'une partie au moins de sa syntaxe".³³

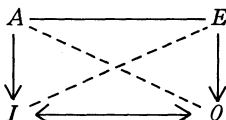
Ce n'est donc pas la réflexivité, mais le surcroît de sa richesse en moyens d'expressions qui fait qu'une langue est contradictoire.

§7. Sur le carré logique des termes: *Vrai, Faux, Dérivable, Réfutable*

On peut avec les termes *Vrai (Analytique, Valide), Faux (Contradictoire,*³⁴ *Contravalide), Démontrable, Réfutable* constituer le carré logique suivant:



Comparons de carré à celui de la Logique classique:



Dans ce dernier carré: *A* et *E* sont des contraires: ils ne peuvent être tous deux vrais, mais peuvent être tous deux faux. On a figuré cette relation par un trait plein. Les couples *A* et *O*, *E* et *I* sont des contradictoires; ils ne peuvent être ensemble ni vrais ni faux. On a figuré cette relation par un trait en pointillé. Les couples *A* et *I*, *E* et *O* sont subalternes: $A \supset I$ et $E \supset O$. On a figuré leurs relations par des flèches simples. Enfin, *I* et *O* sont subcontraires: ils ne peuvent être ensemble faux, mais peuvent être ensemble vrais. On a figuré leur relation par une flèche double.

Les relations, dans le carré des termes sémantiques et syntactiques, sont très différentes. "Vrai" et "Faux" sont des contradictoires: ils ne peuvent être ensemble ni vrais, ni faux. Les couples "Vrai" et "Réfutable", "Faux" et "Démontrable" sont des contraires; ils ne peuvent être vrais ensemble, mais peuvent être faux ensemble. Les couples "Vrai" et "Démontrable", "Faux" et "Réfutable" sont anti-subalternes. En effet: "Démontrable" \supset "Vrai" et "Réfutable" \supset "Faux". On a figuré leurs relations par une flèche de sens ascendant. Enfin, "Démontrable" et

32. Carnap, *op. cit.*, §60 c, p. 218.

33. In Ladrière, *op. cit.*, p. 117.

34. Les concepts sont considérés dans une Langue logique, dont toutes les propositions sont analytiques ou contradictoires.

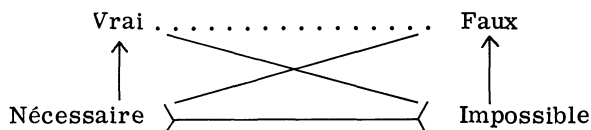
“Réfutable” ne peuvent pas être ensemble vrais et peuvent être faux ensemble; ce sont donc des contraires et comme cette relation est aussi la converse de la subcontrariété, on l’a figurée par une flèche double renversée.

On peut donc déduire les relations formelles du carré logique “syntactique” à partir des relations formelles du carré classique en échangeant les relations de contradiction et celles de contrariété et en renversant le sens des relations de subalternation et de subcontrariété.

Ces changements paraissent fondamentaux dans les théorèmes de limitation. Il y a des indécidables: ce sont les propositions qui ne sont ni démontrables, ni réfutables dans un formalisme donné. Le vrai n’est pas toujours démontrable: on doit donc limiter les moyens d’expression d’un formalisme si l’on veut conserver sa cohérence et, par conséquent, son utilité.

§8. *Sur le carré des termes: Vrai, Faux, Nécessaire, Impossible dans une Langue non-logique et sur le lien entre les antinomies et les théorèmes de limitation.*

Pour une langue non-logique, dans laquelle une proposition vraie peut ne pas être analytique, par exemple dans laquelle il y a des axiomes du genre du Postulat d’Euclide, on obtient avec les prédicats: “Vrai”, “Faux”, “Nécessaire”, “Impossible” un carré semblable au premier de ceux qu’on a examinés:



Il semble que, dans la célèbre querelle philosophique des futurs contingents, on ait souvent raisonné comme si, les propositions vraies l’étant de toute éternité, elles étaient, de ce fait, nécessaires; la difficulté métaphysique concernant la préscience divine et la liberté humaine résulte, en partie, de cette illusion logique, selon laquelle une proposition ne peut être toujours vraie sans l’être nécessairement.

Le carré supposé des termes dans la perspective des partisans de la prédestination et de la nécessité universelle paraît construit de la façon suivante. 1°) Il y a contradiction entre le Vrai et le Faux. 2°) Le Vrai et le Faux étant éternels, ils sont nécessairement ce qu’ils sont; donc le Nécessaire est subalterne du Vrai et l’Impossible est subalterne du Faux. 3°) Le Vrai est aussi subalterne du Nécessaire et l’Impossible du Faux; 4°) donc le Vrai équivaut au Nécessaire et le Faux à l’Impossible; 5°) donc le Nécessaire et l’Impossible sont des contradictoires. Dans ce nouveau carré, toutes les relations entre les termes sont réduites à la contradiction ou à l’équivalence.

C’est évidemment la seconde de ces affirmations qui nous paraît étrange. Vrai, Faux, Nécessaire et Impossible n’ont de sens que relativement à une langue donnée. Or dans cette langue, le Nécessaire et

l'Impossible sont des contraires, si cette langue est non-logique. Leur statut apparaîtra plus clairement au moyen de l'argument suivant, non rigoureux, il est vrai.

Soit **P** cette langue, qu'on a supposée non-logique. Les concepts de Vrai et de Faux sont exprimés dans une méta-langue sémantique **SMP**. Cette méta-langue ne peut être contenue dans **P**, sous peine de rendre possible l'Epiménide. Elle est donc plus riche que **P** et contient le prédicat "Vrai" de **P**, mais, sous peine d'incohérence, ne contient pas son propre prédicat "Vrai". Pour exprimer ce dernier, il faut construire une méta-méta-langue sémantique **SM(SMP)**, plus riche que **SMP**, et ainsi de suite à l'infini. Or la situation n'est pas le même pour le prédicat "Nécessaire" (resp. "Impossible"). Une proposition de **P** est dite nécessaire (resp. impossible), si et seulement si ou bien elle est un axiome (resp. la négation d'un axiome) logique de **P**, ou bien elle est un théorème (resp. un "anti-théorème") logique de **P**, un théorème logique étant un théorème obtenu à partir des axiomes logiques par le seul usage des règles de déduction logique au cas où il existe des règles non-logiques de déduction. Ce sont là des notions syntactiques qu'on peut formuler dans une métalangue syntactique de **P** ou plus exactement de la partie propre logique **L** de **P**. Soit **ML** ce langage. Nous sommes ramenés à un cas particulier des théorèmes de limitation.

On peut reprendre le raisonnement des pages 13-14, en remplaçant, dans le menteur, "faux" par "non-nécessaire", puis par "impossible". Dans aucun de ces cas, une contradiction ne se produit, étant donnée la compatibilité, dans une langue non logique, des propriétés "Vrai" et "Non-Nécessaire" - respectivement "Faux" et "Non-Impossible". La contingence n'est ici que le doublet de l'indécidable. Certaines propositions de **P** ne sont vraies en **SMP** qu'en n'étant ni nécessaires, ni impossibles, comme, dans les métalangues sémantiques de langues formelles contenant l'Arithmétique, il existe une sentence qui n'est vraie que si elle n'est pas un théorème.

La confusion de l'opposition entre Vrai et Faux et de l'opposition entre Nécessaire et Impossible provient de ce qu'on est passé subrepticement d'une Langue non-Logique où les vérités de fait sont distinguées des vérités de raison à une Langue logique où les vérités de fait ne sont de fait qu'en vertu de notre ignorance et sont de raison pour le regard de Dieu. Il est curieux de constater que les partisans de la Nécessité Universelle ont souvent cru défendre et même sauver la science en supposant une Langue parfaite, fermée sur elle-même, complète au point de vue logique et étrangère à toute contingence au point de vue réel, convenable à l'omniscience et à la présience. Or une telle Langue est nécessairement incohérente.

Le rapport des théorèmes de limitation aux antinomies est ainsi lié au genre d'opposition des concepts en vertu desquels une langue est susceptible de donner en elle-même une image d'elle-même. Quand cette opposition est celle de la contradiction, elle produit l'antinomie. Car les contradictoires épuisent pour ainsi dire l'univers du discours et leur

disjonction constitue une tautologie. Les théorèmes de limitation sont, au contraire, associés à des oppositions de contrariété. Les contraires ne dessinent qu'un domaine dans l'univers; ils peuvent être tous deux faux et laissent la place libre pour un troisième terme. Quand l'opposition de contradiction porte sur l'image de la langue dans la langue, elle la ferme sur elle-même; de cette fermeture naît l'antinomie, qui disparaît lorsque l'opposition a lieu dans une métalangue différente de la langue-objet. Au contraire, l'opposition de contrariété peut se produire dans une langue-objet cohérente et néanmoins la réfléchir, puisqu'elle n'embrasse pas l'univers du discours et il est alors possible qu'elle donne lieu aux théorèmes de limitation.

Collège de France
Paris, France